P4.11 AN EVALUATION OF NSSL'S NEAR-STORM ENVIRONMENT ALGORITHM

Shannon A. Myers* and Travis M. Smith
CIMMS/Univ. of Oklahoma and NOAA/NSSL, Norman, Oklahoma

1. INTRODUCTION

The National Severe Storms Laboratory (NSSL) uses a wide variety of radar-based algorithms to aid forecasters with the assessment of severe weather conditions. Many of these algorithms, such as the Mesocyclone Detection Algorithm and the Hail Detection Algorithm, are further augmented with environmental data. These environmental data are provided to these algorithms by the Near-Storm Environment (NSE) algorithm (Lee et al. 1998). Currently, the NSE algorithm ingests forecast fields from NCEP's Rapid Update Cycle (RUC; Benjamin et al, 2002) model and produces more than 125 2-D and 3-D grids of derived environmental data, many of which are not standard output for the RUC.

When running in real-time, there is usually a delay of 45 minutes to 1 hour between the valid time of the RUC model analysis and time that the model data are actually received by users. Thus, the most recent data processed by the NSE algorithm and provided to other algorithms may be as much as 2 hours old. In order to provide more up-to-date information to NSSL's severe storm analysis algorithms, we have initiated a study to compare techniques that are being considered to insert more recent data into the NSE algorithm than is provided by the RUC analysis alone. Since other algorithms that operate in regions of convection use the NSE algorithm's output, we focus on only those specific grid boxes for which cloud-to-ground lightning strikes have been detected. This study uses statistical analysis to compare environmental fields produced by the NSE algorithm when different ambient fields are provided to the algorithm. This manuscript describes the data analysis technique and some preliminary results. Detailed results will be presented both at the conference and online.

2. DATA

The data used for the evaluation are collected at NSSL daily. The 20km RUC data is obtained directly from NCEP, and lightning data used to determine areas of convection are obtained from the National Lightning Detection Network. Surface data are obtained from METAR stations and mapped to a grid using a 2-pass Barnes objective analysis.

2.1 *Data Collection*

These preliminary results use model, surface, and lightning data from the month of April 2002. 83% of the month of April (600 hours of data) is present in the current data set. Further work is planned to include more days in the final data set. The NSE algorithm provides environmental parameters based on input from the RUC model. RUC grid points used in the survey were identified based on convection present in the grid, identified by 5 or more lightning strikes in the grid box during the hour following the analysis time. There were 52889 grid points that met these criteria that were used in the survey.

2.2 *Data Considerations*

The RUC model output used by the NSE algorithm includes an analysis and a 1-hour forecast. The RUC analysis is available nearly one hour after the valid time. This results in one-hour old environmental data being ingested by the real-time NSE algorithm. To create a more accurate depiction of the ambient environment near storms, the most recent surface data should be input into the NSE algorithm. A Barnes objective analysis scheme (Barnes 1964) is used to replace the RUC surface data fields with the current hour surface observations. These are combined with the RUC upper air information from the analysis time or the 1-hour forecast time. The RUC analysis or 1-hour forecast from the previous hour is used as first guess in the objective analysis scheme. These two methods are tested against the current method of persistence (RUC analysis from one hour ago). The RUC analysis valid at the time of the surface observations was used as validation for this study. The study hopes to attain an improvement in the accuracy of the NSE algorithm parameters by giving the algorithm more current surface information.

The three methods compared in the study are shown in Table 1. The first column is the method number. The second column describes the origin of the surface data, and the time that the surface data is originating from. The third column describes the origin of the upper-air data, and the time of data obtained from the RUC.

Method	Surface / time	UA / model time
1	RUC T-1	analysis T-1
2	METAR T=0	analysis T-1
3	METAR T=0	1hr forecast T-1

Table 1: Environmental data methods used as input to the NSE algorithm. "T-1" represents the previous hour's data, while "T=0" represents the current hour.

* *Corresponding author address*: Shannon A. Myers, National Severe Storms Laboratory, 1313 Halley Circle, Norman, Oklahoma, 73069. Email: Shannon.Myers@noaa.gov.

Again, these procedures give us three separate forecasts that can be used in the NSE algorithm:

1. Persistence (the previous hour RUC analysis)
2. RUC upper-air analysis from one hour ago merged with current hour surface data
3. RUC 1-hour upper-air forecast from one hour ago merged with current surface data

Each method has its strengths and weaknesses. The first method provides the NSE algorithm with surface and upper air data grids that are from the same initialization time, therefore preventing data discontinuities that may occur when new surface 2-D analysis grids are inserted into the same domain as upper-air grids from a different time. The drawback to method 1 is that the most recent data may be up to two hours old. The passage of fronts or actual convection initiated during the hour may not be reflected in the data. Method 2 has the advantage of the most recent surface data being added to an upper air analysis that is one hour old. The surface data, in this case, will be only minutes to one hour old when it is being examined as the "latest available", while the upper air data is still one to two hours old. This partially solves the problem of frontal or other boundary passage and surface temperature change, but introduces a potential data discontinuity. Method three also uses current hour surface data, but with the one-hour upper air forecast valid for the current hour. Although this method may provide a more accurate upper air forecast, convection initiated by the RUC within this first hour of its forecast run could affect these ambient fields.

These represent three possible methods to provide the NSE algorithm with convective environmental parameters. In the study, RUC grid points were identified as being in regions of convection when five or more lightning strikes occurred in a grid box in the hour following the analysis time. This insures the environmental variables used in the study are from a preconvective environment. Grid points with consecutive hours of convection will be evaluated separately in a future study.

Once the NSE algorithm is run using the three different fields, grid points with convection are identified, and different parameters may be examined for each of the methods. Table 2 lists a small portion of these variables and their abbreviations.

3. ANALYSIS

A common method for comparing forecasted variables versus observed variables is the mean squared error (MSE). As observed by Murphy and Winkler (1986), MSE is a valuable method for evaluating the accuracy of forecasts. One of the valuable properties of MSE is that it can be decomposed into other useful statistical variables. The formula used for MSE is as follows:

$$MSE = Var(f) + Var(x) - 2Cov(f,x) + [E(f) - E(x)]^2$$

PMSL	Mean sea level pressure
H273	Height of the 273 K isotherm
UCAP	Convective available potential energy of most unstable parcel in lowest 300mb
ULCL	Height of the LCL based on most unstable parcel
UEHI	Energy-helicity index for the most unstable parcel in lowest 300mb
SCAP	Convective available potential energy of surface parcel
SEHI	Energy-helicity index for the surface parcel
MCNV	Surface moisture convergence
LAPS	Mean lapse rate in the 850-500 mb layer
DDCU	Downdraft convective available potential energy
SRH3	Estimated 0-3 km storm relative helicity
06SM	0-6 km shear magnitude
ACAP	Convective available potential energy averaged over the lowest 100mb
HWBZ	Height (m) of the wet-bulb=273K level
TMPF	Temperature (F)
DWPF	Dew point (F)

Table 2: Environmental parameter abbreviations and their definitions.

MSE values for NSE parameters			
	Persistence	Method 2	Method 3
PMSL	0.704397	0.721658	0.831801
H273	9435.05	9435.05	5843.758
UCAP	239175.6	284705.4	210663.3
ULCL	157814.9	163043.7	129604.4
UEHI	0.715583	1.020788	0.702101
SCAP	286878	473605	414978
SEHI	0.510008	0.683246	0.587175
MCNV	7.911052	9.568104	9.395507
LAPS	0.06108	0.06108	0.03455
DDCU	27043.11	26900.80	18622.99
SRH3	4420.271	7468.717	5279.302
06SM	27.4639	50.5982	44.4094
ACAP	85699	130850	108976
HWBZ	10470.73	10470.73	3134.98
TMPF	6.57131	5.09003	3.88768
DWPF	13.56335	13.61491	4.65519

Table 3: Mean squared error values for NSE parameters shown for each of the three forecast methods.

where *f* represents the forecasted values and *x* represents the observed values. *Var(f)* and *Var(x)* are the variance of the forecasts and observations, respectively. *Cov(f,x)* is the covariance, and *E(f)* and *E(x)* are the expected values of the forecasts and observations, where *E(f)* – *E(x)* is the bias. The MSE for each parameter is shown in Table 3. The highlighted values represent the method with the lowest MSE for each parameter, indicating the best method to use for the particular variable.

For this early part of our study, one method did not clearly excel above the others. In many cases persistence (method 1) is the best approximation for the preconvective environment, and in other cases the RUC 1-hour forecast merged with the current surface data (method 3) was the best approximation.

As an example, the decomposition of MSE for height of the wet-bulb zero isotherm (HWBZ), is seen in Table 4. Method 3 is the best approximation for this parameter, as it is much lower than for the other two methods. The variance (*Var(f)*) and covariance (*Cov(f,x)*) for Method 3 is smaller than for the other two methods, and the bias is also about 7 m smaller. The small negative bias and the lower variance values contribute to the smaller MSE for method 3.

	Observed	**Persistence**	**Method 2**	**Method 3**
Mean	3344.4	3332.9	3332.9	3339.6
MSE		10470	10470	3134
MAE		73.92	73.92	37.68
Var	123736	127922	127922	121358
Bias		-11.57	-11.57	-4.79
Cov		120660	120660	120991

Table 4: MSE decomposition for height of the wet bulb zero (HWBZ; meters). MAE is the Mean Absolute Error.

4. CONCLUSIONS

Of the three methods used to calculate environmental parameters, it is not clear that one is dominant based on our small initial data set. The MSE for method 2 was never lower than that of persistence or of method 3. It may be that persistence works best for surface data and the one-hour forecast is best suited for upper-air. The results of our study are very preliminary. The addition of more data and further analysis of the data sets may yield clearer results. No absolute conclusions may be drawn yet as to which forecast method prevails.

5. FUTURE WORK

More data are being added to the data set, and will be evaluated in a similar manner to data presented here. By increasing the sample size, one method may be seen as superior to the others. A dynamic webpage is planned, and all data will be presented graphically. More analysis and results will be featured on the webpage, at http://www.nssl.noaa.gov/nse. Future studies are planned to evaluate the data by geographic regions, as results may differ in between relatively flat and mountainous terrain. It may be that different methods are more appropriate depending on the geographic region. Additionally, confidence intervals for MSE will be calculated. With a larger data set, results about which method is most appropriate for determining environmental fields for input into other algorithms may become more obvious.

6. ACKNOWLEDGEMENTS

The authors would like to thank Phillip Spencer and Greg Stumpf for reviewing this manuscript.

7. REFERENCES

Benjamin, S.G., J.M. Brown, K.J. Brundage, D. Devenyi, G.A. Grell, D. Kim, B.E. Schwartz, T.G. Smirnova, T.L. Smith, S. Weygandt, and G.S. Manikin, 2002: RUC20 - The 20-km version of the Rapid Update Cycle NWS Technical Procedures Bulletin No. 490.

Barnes, S.L., 1964: A Technique for Maximizing Details in Numerical Weather Map Analysis. *J. Appl. Meteor.,* **3**, 396-409.

Lee, R. L., G.J. Stumpf, and P.L. Spencer, 1998: Should geographic region or near-storm environment dictate. WSR-88D algorithm adaptable parameter settings. Preprints, *19th Conference on Severe Local Storms*, Minneapolis, Amer. Meteor. Soc.,784-787.

Murphy, A.H., and R.L. Winkler, 1986: A General Framework for Forecast Verification. *Mo. Weather Rev.,* **115**, 1330-1338.